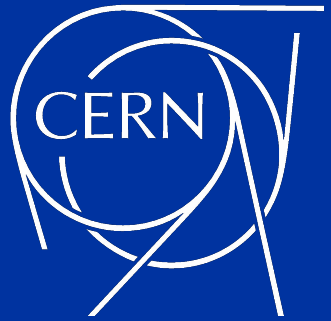

Python for HPC

Tools to boost your everyday computing

Kostis Iliakis (konstantinos.iliakis@cern.ch)

13/10/2022

Contents

How to access and run the notebooks

- Github repository: https://github.com/kiliakis/Py4HPC
- Notebooks 01, 02, 03: Can run in Swan or your PC.
- Notebooks 04, 05: Need a GPU
  - ml.cern.ch (access only from within CERN)
  - Google colab: https://colab.research.google.com/
  - Or your PC if it has an Nvidia GPU
- **You do not need to run the notebooks during this tutorial!**
  - You may do it later if interested.

Notebooks presentation..

Summary

- **Profiling is essential for performance optimization**

Tool	Notebook	Command Line	Profiling
`timeit`	Yes	No	Single function
`line_profiler`	Yes	Yes	Line by line
`cProfile`	Yes	Yes	Call stack

- **Performance Optimization/ Acceleration Tools**

Tool	CPU	Multicore	GPU (Nvidia)	GPU (AMD)	User functions	Numpy/ Scipy
`Numba`	Yes	Yes	Yes	No	Yes	No
`CuPy`	No	No	Yes	Yes	Yes	Yes

Other HPC tools

- **Dask** (https://www.dask.org/): Scale calculations in multiple nodes, supports libraries like NumPy, Pandas, scikit-learn.
- **MPI4PY** (https://mpi4py.readthedocs.io/en/stable/): Interface to the MPI library in Python, for distributed computing.
- **Cython** (https://cython.org/): A superset of Python, that makes writing C extensions for Python easy.